REMARKS

Claims 1-70 are pending. Claims 45-70 have been withdrawn from consideration. By the Amendment, Claims 1-5, 8, 10-13, 16-18, 20-27, 30-32, 40-41 and 44 are amended.

In the Office Action, the Examiner rejects Claims 1-9 under 35 U.S.C. § 102(e) over U.S. Publication No. 2003/0126056 A1 to Hausman, *et al.* (Hausman). The Examiner also rejects Claims 1-4, 13-15 and 17-21 under 35 U.S.C. § 102(e) over U.S. Patent No. 6,615,312 to Hamlin, *et al.* (Hamlin). The Examiner also rejects Claims 8, 10-12 and 16 under 35 U.S.C. § 103(a) over Hamlin in combination with U.S. Publication No. 2003/0018661 A1 to Darugar (Darugar). The Examiner also rejects Claims 22-32 and 40 under 35 U.S.C. § 103(a) over Hamlin in combination with U.S. Patent No. 5,785,360 to Zbikowski, *et al.*(Zbikowski). The Examiner also rejects Claims 33-39 and 41-44 under 35 U.S.C. § 103(a) over Hamlin in combination with Zbikowski and Darugar.

These rejections are respectfully traversed.

Prior to discussion, the Examiner may find the following definitions useful to establish common understanding.

Source data: A collection of items of data, which can for example be provided as input to a computer program in any kind of readable storage or transmission media, file, or stream, which include individual items. The individual items can include or comprise, for example, a recognizable single fact or business measurement. Examples of source data include: a spreadsheet or database table; a query resulting in data extracted from a database table; a comma-separated-variables file; an XML or HTML file or stream; a data stream output from a computer

to one or more of a display screen, a memory, a hard drive, a CD ROM drive, a floppy disk drive a printer, or other device; and a table of data in a Microsoft Word document.

The collection of data and first metadata recited in Claim 1 can for example be source data.

Metadata: Data about data, for example that defines or characterizes data (e.g., by classifying items of source data). Metadata can include documentation or information describing characteristics, such as name; size, attributes, numeric or string constraints, conditions, optionality, and so forth. Metadata can include or indicate relationships with data or interrelationships among data, and metadata can be multidimensional. Classification metadata, for example, is often presented to computer programs in the form of a schema, data model, taxonomy, or dictionary. Contextual metadata may specify information about the data item being described, such as the reporting period, entity (business, government department, individual, etc.) that data item describes, and the reporting scenario; measurement metadata may specify the unit of measure of a data item (feet or meters, dollars or yen). Interrelationship metadata (which can be considered a form of contextual metadata) may organize or group data items for the same employee such as name, address, and department numbers together; footnote metadata may interrelate multiple data items with the same footnote reference, and can be considered a form of contextual metadata.

Mapping: To make a correspondence between an item of source data and any dimension of metadata; for example by associating a data item with a classifying description provided in metadata, such as to classify or identify the items of source

data; associating a data item with contextual metadata (such as what date or entity it might apply for); associating a data item with a unit of measure (euro, yen); interrelating data items (this name and address are of one employee); and interrelating data items that have a common footnote reference. Some examples of mapping are: by position, placement or location cues, for example of information within the collection of data and/or with respect to other data or metadata within the collection, as in a spreadsheet which has columns identified for "name", "address", "earnings", as in balance sheets with chart of account labels to the left of numbers and underlines above totals, or where for example a department code is in one location, and a reporting period is in another location. Logical expressions can be used to help to identify data, for example to process positional cues or other information, for example a department code in one location, and a reporting period in another location. Mapping can also be by syntax (such as XML with tags such as <name>Joe Smith</name>), or database query.

As outlined in the background section of the present application, currently there are thousands upon thousands of software programs installed in millions of computers that cannot transfer data to each other without extensive and expensive custom programming. For example, large companies with many branches or subsidiaries often find that the accounting or operating software programs used by one division or subsidiary are not compatible with the software used by other divisions or subsidiaries or the central corporate programs. This requires substantial conversions of data and often results in a great deal of data reentry along with the costs and data integrity problems that attend data entry.

Because of the great variety of programs, operating systems and software standards currently used by software developers there is a great deal of incompatibility between suppliers and their customers. This also requires substantial conversions of data and often results in a great deal of data reentry with corresponding costs and opportunities to introduce errors into the data. The unstructured and undefined nature of the current computer software environment imposes great burdens and expense on regulatory organizations such as the Federal Financial Institutions Examination Council, the Federal Securities and Exchange Commission, Federal and State tax authorities, banks, etc. and the companies reporting to them.

To overcome this problem many standards organizations have been formed to establish defined input/output vocabularies for use with the XML (eXtensible Markup Language) file format. The emergence of new standards for defining metadata such as eXtensible Business Reporting Language, or XBRL, provide an agreement of syntax for transfer, and the emergence of standard metadata (such as standards organization produced schemas and taxonomies) provide an agreement of semantics for such transfer.

However, it is a challenge to automatically or semi-automatically convert conventional documents or data into outputs tagged with the standardized metadata or Information Labels called for by XML, XBRL, or other standards committees. This challenge may be compounded when companies take years to migrate to new software products that are designed to output the appropriate Information Labels, and in some cases where it is virtually impossible to replace legacy software systems, this may never happen.

Exemplary embodiments of the present invention solve the problem of semantic attribution of data items, by identifying data in a conventional document and adding metadata that will allow or enhance semantic processing. For example, exemplary embodiments can add metadata to a conventional document to produce an XBRL-compliant document whose data can be easily understood and processed by a large variety of computer platforms.

In accordance with exemplary embodiments, metadata in a file of source data or in a captured data stream, **together with location of information within the file**, for example in relationship to other data, can be used to identify data in the file. See, for example, numbered paragraphs [0029] and [0039] of the present application. Then, additional metadata referring to the identified data can be added to the file or captured data stream. The added metadata can then allow the data to be easily consumed by a broad variety of computing platforms.

Turning now to the references cited by the Examiner, Applicants note that Hamlin discloses a method for processing file system service requests so that data read to or from a hard disk drive can be done without changing the mode of the disk drive, regardless of whether the data is stream data or non-stream data. Hamlin is directed to disk controller hardware and operating systems buffering techniques - Hamlin addresses disks, files, stream and non-stream data, and is at the level of computer disk hardware, controllers, operating systems, drivers, and language processor read/write operations.

Hamlin fails to disclose or suggest software processing of data objects inside streams or collections of data, so that the data can be labeled (e.g., with metadata) to enhance semantic processing of the data.

In particular, Hamlin discloses receiving a file system service request, recording whether the request is for non-stream data or stream data, associating disk locations of the disk drive with the request, and then preparing a command to the disk drive that includes an indication whether the request is for non-stream data or stream data. See, for example, Hamlin's abstract and Figure 1. The description of Figures 2-3 which is cited by the Examiner and found at column 12, lines 43-67, further discloses that in response to the request, the file name of the requested data is correlated with logical spaces (e.g. block addresses) on the disk drive. The file system can keep track of which data are streaming data, for example via a record 206 that indicates which logical block addresses of the disk drive correspond to or contain stream data. Thus the record 206 can be used to alert the disk drive when stream data are involved in the request. See for example column 12, lines 59-65. The system service request can also indicate whether the data is stream data (see e.g. column 12, lines 57-58). Thus the disclosure of Hamlin is not relevant to processing data to correlate items of data with their semantic meaning as described by metadata.

Accordingly, Hamlin fails to disclose or suggest a computer implemented method for adding metadata to a collection of data and first metadata wherein the first metadata are associated with the data, including *identifying data in the collection based on the first metadata **and** one or more locations of the data and/or the first metadata in the collection, and adding second metadata to the collection based on the identified data* as recited in Claim 1.

Zbikowski discloses different ways of storing data and metadata in a file system, for example "onodes" that are groups of data streams related by

functionality. See, for example, column 6, lines 8-13. As does Hamlin, Zbikowsi discusses operating system and buffering techniques relating to storage of data streams on disk drives. However, like Hamlin, Zbikowski is not relevant to the issues of data item classification and identification by metadata descriptions.

Like Hamlin, Zbikowski fails to disclose or suggest a computer implemented method for adding metadata to a collection of data and first metadata wherein the first metadata are associated with the data, including *identifying data in the collection based on the first metadata **and** one or more locations of the data and/or the first metadata in the collection, and adding second metadata to the collection based on the identified data* as recited in Claim 1.

Hausman and Darugar likewise fail to overcome the deficiencies of Hamlin and Zbikowski.

Hausman is directed to filtering data from a stream by identifying the data using proprietary tags, but fails to disclose or suggest how such tags or metadata are added to the stream in the first place. Hausman addresses a stream as a distribution mechanism of multicast data, such as sports scores or news information, where multiple users filter, extract, and otherwise process data as needed. In contrast, exemplary embodiments of the present application that are encompassed by the claims, map data items into or between data descriptions within metadata regardless of their sources (which could be spreadsheets, data bases, or even sports scores streams such as Hausman addresses). Hausman does not disclose or describe taxonometric or schematic mapping of data, only its distribution.

Darugar discloses a tool to aid a user in mapping elements from a first XML format into a second XML format by associating each element in the first format with

one or more elements in the second format. However, Darugar addresses tag transformation issues, not data identification by metadata description or attribution, as in exemplary embodiments of the present invention.

Thus Hausman and Darugar, when considered both separately and in combination with Hamlin and Zbikowski, a computer implemented method for adding metadata to a collection of data and first metadata wherein the first metadata are associated with the data, including *identifying data in the collection based on the first metadata and one or more locations of the data and/or the first metadata in the collection, and adding second metadata to the collection based on the identified data* as recited in Claim 1.

For at least the above reasons, Applicants respectfully submit that Hamlin, Zbikowski, Hausman and Darugar, when considered both separately and in combination, fail to disclose or suggest the combination of features recited in Claim 1. Claims 2-44 depend variously from Claim 1, and are likewise allowable for at least the same reasons. Withdrawal of the rejections under 35 U.S.C. § 102(e) over each of Hausman and Hamlin, as well as of the rejections under 35 U.S.C. § 103(a) over combinations of Hamlin, Zbikowski and Darugar, is respectfully requested.

Applicants respectfully submit that the application is in condition for allowance. Favorable consideration on the merits and prompt allowance are respectfully requested. In the event any questions arise regarding this communication or the application in general, the Examiner is invited to contact Applicants' undersigned representative at the telephone number listed below.

Respectfully submitted,

BURNS, DOANE, SWECKER & MATHIS, L.L.P.

Date: __August 2, 2004__   By: _____

M. David Ream
Registration No. 35,333

P.O. Box 1404
Alexandria, Virginia 22313-1404
(703) 836-6620